



Munich Personal RePEc Archive

# The revelation principle fails when the format of each agent's strategy is an action

Wu, Haoyang

Wan-Dou-Miao Research Lab

28 November 2020

Online at <https://mpra.ub.uni-muenchen.de/104417/>  
MPRA Paper No. 104417, posted 05 Dec 2020 07:03 UTC

# The revelation principle fails when the format of each agent's strategy is an action

Haoyang Wu

*Wan-Dou-Miao Research Lab, Shanghai, China.*

---

## Abstract

In mechanism design theory, a designer would like to implement a social choice function which specifies her favorite outcome for each possible profile of agents' private types. The revelation principle asserts that if a social choice function can be implemented by a mechanism in equilibrium, then there exists a direct mechanism that can truthfully implement it.

This paper aims to propose a failure of the revelation principle. We point out that in any game the format of each agent's strategy is either an informational message or a realistic action, and the action format is very common in many practical cases. The main result is that: For any given social choice function, if the mechanism which implements it has action-format strategies, then "*honest and obedient*" will not be the equilibrium of the corresponding direct mechanism. Consequently, the revelation principle fails when the format of each agent's strategy is an action.

*Key words:* Mechanism design; Revelation principle.

---

## 1 Introduction

In the framework of mechanism design theory [1–3], there are one designer and some agents labeled as  $1, \dots, I$ .<sup>1</sup> Suppose that the designer would like to implement a social choice function which specifies her favorite outcome for each possible profile of agents' types, and each agent's type is modeled as his privacy. In order to implement a social choice function in equilibrium, the designer constructs a mechanism which specifies each agent's feasible strategy

---

*Email address:* 18621753457@163.com (Haoyang Wu).

<sup>1</sup> In this paper, the designer is always denoted as "She", and the agent is denoted as "He".

set (*i.e.*, the allowed actions of each agent) and an outcome function (*i.e.*, a rule for how agents' actions get turned into a social choice).

The revelation principle is an important theorem in mechanism design theory. It asserts that if a social choice function can be implemented by a mechanism in equilibrium, then it is truthfully implementable. So far, there have been several criticisms on the revelation principle: Bester and Strausz [4] pointed out that the revelation principle may fail because of imperfect commitment; Epstein and Peters [5] proposed that the revelation principle fails in situations where several mechanism designers compete against each other. Kephart and Conitzer [6] proposed that when reporting truthfully is costless and misreporting is costly, the revelation principle can fail to hold.

Different from these criticisms, this paper aims to propose another failure of the revelation principle. The paper is organized as follows: Section 2 analyses two formats of strategy and points out that the action format is very common in many practical cases. Section 3 proposes the main result, *i.e.*, the revelation principle fails when each agent's strategy is action-format. Section 4 draws conclusions. Notations about mechanism design theory are given in Appendix, which are cited from Ref [1].

## 2 Two formats of strategy

**Note 1:** In any game, the format of each agent's strategy is either an informational message or a realistic action.  $\square$

Although the note looks naive, it is not trivial. The reason why we emphasize the distinction of two formats of strategy is that the revelation principle does not hold for the case of action-format strategies, as will be discussed in Section 3. For simplification, in the following discussions we simply assume that in any game all agents' strategies are of the same format, *i.e.*, we omit the case in which some agents' strategies are message-format and other agents' strategies are action-format. Next, we will deeply investigate the two formats of strategy respectively.

### 2.1 Case 1: Mechanism with message-format strategies

**Definition 1:** A message-format strategy of an agent in a mechanism is a strategy represented by an informational message. The information itself contained in the message is just the agent's strategy, which does not need to be carried out realistically in the mechanism.

Practically, only in some restricted cases can each agent's strategy be described as pure information and represented by an informational message. For example, let us consider a chess game, then each player's strategy is message-format, since it is a strategic plan about how to play chess. Similarly, the strategy in a war simulation game is also message-format, since it contains military *plans* of players. However, in many practical cases each agent's strategy cannot be described as pure information but must be described as a realistic action. For example, the strategy in a real war is action-format, since it contains military *actions* of armies.

**Definition 2:** Given a social choice function  $f$ , suppose a mechanism  $\Gamma = (S_1, \dots, S_I, g(\cdot))$  implements it in equilibrium with message-format strategies. To clearly describe the case of message-format strategies, we denote each strategy set  $S_i$  as  $M_i$ , and each agent  $i$ 's strategy function as  $m_i(\cdot) : \Theta_i \rightarrow M_i$ , where  $\Theta_i$  is agent  $i$ 's type set. The outcome function  $g(\cdot)$  is denoted as  $g_m(\cdot) : M_1 \times \dots \times M_I \rightarrow X$ , where the input parameters are message-format strategies and  $X$  is the set of outcomes. Hence, the mechanism  $\Gamma$  is denoted as  $\Gamma_m = (M_1, \dots, M_I, g_m(\cdot))$ . The game induced by  $\Gamma_m$  is denoted as  $G_m$ , which works in a one-stage manner:

*Step 1:* By using the strategy function  $m_i(\cdot)$ , each agent  $i$  with private type  $\theta_i$  sends a message  $m_i(\theta_i)$  to the designer.<sup>2</sup>

*Step 2:* The mechanism  $\Gamma_m$  yields the outcome  $g_m(m_1(\theta_1), \dots, m_I(\theta_I))$ . Here, each agent  $i$ 's utility is denoted as  $u_i(g_m(m_1(\theta_1), \dots, m_I(\theta_I)), \theta_i)$ .

**Definition 3:** Suppose the game  $G_m$  has a Bayesian Nash equilibrium, denoted as  $m^*(\cdot) = (m_1^*(\cdot), \dots, m_I^*(\cdot))$ , i.e., for all  $i$  and all  $\theta_i \in \Theta_i$ ,  $\hat{m}_i \in M_i$ ,

$$E_{\theta_{-i}}[u_i(g_m(m_i^*(\theta_i), m_{-i}^*(\theta_{-i})), \theta_i) | \theta_i] \geq E_{\theta_{-i}}[u_i(g_m(\hat{m}_i, m_{-i}^*(\theta_{-i})), \theta_i) | \theta_i].$$

Consider this equilibrium, there is a compound mapping from agents' possible types  $\hat{\theta} = (\hat{\theta}_1, \dots, \hat{\theta}_I) \in \Theta$  into the outcome  $g_m(m^*(\hat{\theta}))$ , which is equal to  $f(\hat{\theta})$ . Based on the compound mapping, we construct a direct mechanism  $\bar{\Gamma}_m = (\Theta_1, \dots, \Theta_I, g_m(m^*(\cdot)))$ .<sup>3</sup>

**Definition 4:** The direct mechanism  $\bar{\Gamma}_m$  induces a *one-stage* direct game  $\bar{G}_m$ , which works as follows:

*Step 1:* Each agent  $i$  with private type  $\theta_i$  individually reports a type  $\hat{\theta}_i \in \Theta_i$

<sup>2</sup> In the following discussions, we denote each agent  $i$ 's true type as  $\theta_i$ , and his any possible type as  $\hat{\theta}_i \in \Theta_i$ .

<sup>3</sup> Although  $g_m(m^*(\hat{\theta})) = f(\hat{\theta})$  for any  $\hat{\theta} = (\hat{\theta}_1, \dots, \hat{\theta}_I) \in \Theta$ , the outcome function of the constructed direct mechanism  $\bar{\Gamma}_m$  must be the compound mapping  $g_m(m^*(\cdot))$  instead of  $f(\cdot)$ . The reason is straightforward: if the outcome function of  $\bar{\Gamma}_m$  is simply written as  $f(\cdot)$ , then  $\bar{\Gamma}_m$  will become a naive direct mechanism  $(\Theta_1, \dots, \Theta_I, f(\cdot))$ , which is irrelevant to the original mechanism  $\Gamma_m = (M_1, \dots, M_I, g_m(\cdot))$ , and indeed cannot implement  $f(\cdot)$  at all.

to the designer. Here, each agent  $i$  does not need to be “honest”, *i.e.*,  $\hat{\theta}_i$  can be different from agent  $i$ ’s private type  $\theta_i$ .

*Step 2:* By using the equilibrium strategy functions  $m^*(\cdot) = (m_1^*(\cdot), \dots, m_I^*(\cdot))$ , the direct mechanism  $\bar{\Gamma}_m$  calculates  $m^*(\hat{\theta}) = (m_1^*(\hat{\theta}_1), \dots, m_I^*(\hat{\theta}_I))$ , and then yields the outcome  $g_m(m^*(\hat{\theta}))$ .

**Note 2:** Obviously,  $m^*(\hat{\theta}) = (m_1^*(\hat{\theta}_1), \dots, m_I^*(\hat{\theta}_I))$  are pure information, and hence are message-format. Actually, only when each agent  $i$ ’s strategy  $m_i(\cdot)$  is message-format can the calculated results  $m^*(\hat{\theta})$  be *legal* parameters of the outcome function  $g_m(\cdot)$ . As a comparison, when each agent’s strategy is action-format, then any calculation about it will only be a message-format description, but not a legal action-format strategy.  $\square$

**Note 3:** By Definition 3, since  $m^*(\cdot) = (m_1^*(\cdot), \dots, m_I^*(\cdot))$  is the equilibrium of the game  $G_m$ , then  $m_i^*(\theta_i)$  is the optimal choice for each agent  $i$  given that all other agents send  $m_{-i}^*(\theta_{-i})$ . Therefore, in the direct game  $\bar{G}_m$ , each agent  $i$  will find truth-telling  $\hat{\theta}_i = \theta_i$  to be the optimal choice given that the others agents tell the truth  $\hat{\theta}_{-i} = \theta_{-i}$ , and the final outcome will be  $g_m(m^*(\theta))$ , which is equal to  $f(\theta)$  for all  $\theta \in \Theta$ . Thus, for the case of message-format strategies, truth-telling is a Bayesian Nash equilibrium of the direct game  $\bar{G}_m$ . Consequently, *the revelation principle holds when each agent’s strategy is message-format*.  $\square$

## 2.2 Case 2: Mechanism with action-format strategies

**Definition 5:** An action-format strategy of an agent in a mechanism is a strategy represented by a realistic action, which should be performed by himself practically.

For example, let us consider a tennis game, then each player’s strategy is his realistic action of playing tennis, but not any informational plan of how to play tennis. Another interesting example is the auction. At first sight, each bidder’s bid is pure information and looks like a message-format strategy. However, in many practical cases, only the bid information itself is not enough to be a full strategy of the auction. Besides announcing a message-format bid, each bidder’s strategy should include performing a *realistic* action (*e.g.*, paying his bid to the auctioneer) if he wins the auction. Hence, in many practical cases, an auction is indeed a game with action-format strategies.

**Definition 6:** Given a social choice function  $f$ , suppose a mechanism  $\Gamma = (S_1, \dots, S_I, g(\cdot))$  implements it in equilibrium with action-format strategies. To clearly describe the case of action-format strategies, we denote each strategy set  $S_i$  as  $A_i$ , and each agent  $i$ ’s strategy function as  $a_i(\cdot) : \Theta_i \rightarrow A_i$ . The

outcome function  $g(\cdot)$  is denoted as  $g_a(\cdot) : A_1 \times \cdots \times A_I \rightarrow X$ , where the input parameters are action-format strategies. Hence, the mechanism  $\Gamma$  is denoted as  $\Gamma_a = (A_1, \cdots, A_I, g_a(\cdot))$ . The game induced by  $\Gamma_a$  is denoted as  $G_a$ , which works in a one-stage manner:

*Step 1:* By using action-format strategy functions  $(a_1(\cdot), \cdots, a_I(\cdot))$ , agents  $1, \cdots, I$  with private types  $(\theta_1, \cdots, \theta_I)$  perform the action-format strategies  $(a_1(\theta_1), \cdots, a_I(\theta_I))$ .<sup>4</sup>

*Step 2:* The mechanism  $\Gamma_a$  yields the outcome  $g_a(a_1(\theta_1), \cdots, a_I(\theta_I))$ .<sup>5</sup> Here, each agent  $i$ 's utility is denoted as  $u_i(g_a(a_1(\theta_1), \cdots, a_I(\theta_I)), \theta_i)$ .

**Definition 7:** Suppose the game  $G_a$  has an equilibrium  $a^*(\cdot) = (a_1^*(\cdot), \cdots, a_I^*(\cdot))$ , i.e., for all  $i$  and all  $\theta_i \in \Theta_i$ ,  $\hat{a}_i \in A_i$ ,

$$E_{\theta_{-i}}[u_i(g_a(a_i^*(\theta_i), a_{-i}^*(\theta_{-i})), \theta_i) | \theta_i] \geq E_{\theta_{-i}}[u_i(g_a(\hat{a}_i, a_{-i}^*(\theta_{-i})), \theta_i) | \theta_i].$$

Consider this equilibrium, there is a compound mapping from agents' possible types  $\hat{\theta} = (\hat{\theta}_1, \cdots, \hat{\theta}_I) \in \Theta$  into the outcome  $g_a(a^*(\hat{\theta}))$ , which is equal to  $f(\hat{\theta})$ . Based on the compound mapping, we construct a direct mechanism  $\bar{\Gamma}_a = (\Theta_1, \cdots, \Theta_I, g_a(a^*(\cdot)))$ .

**Definition 8:** According to Myerson [2], the direct mechanism  $\bar{\Gamma}_a$  induces a *multistage* direct game  $\bar{G}_a$ , which works as follows:

*Step 1:* Each agent  $i$  with private type  $\theta_i$  individually reports a type  $\hat{\theta}_i \in \Theta_i$ . Here each agent does not need to be “*honest*”, i.e.,  $\hat{\theta}_i$  can be different from  $\theta_i$ .

*Step 2:* The designer returns a suggestion to each agent  $i$ , here the suggestion is just the message-format description of action  $a_i^*(\hat{\theta}_i) \in A_i$ . In order to specify the suggestion's format more clearly, we denote the suggestion as  $a_i^m(\hat{\theta}_i)$ ;

*Step 3:* Each agent  $i$  individually performs an action  $\hat{a}_i \in A_i$ . Here each agent  $i$  does not need to be “*obedient*”, i.e.,  $\hat{a}_i$  can be different from  $a_i^m(\hat{\theta}_i)$ .

*Step 4:* After observing that all actions  $\hat{a}_1, \cdots, \hat{a}_I$  have been performed, the direct mechanism  $\bar{\Gamma}_a$  yields the outcome  $g_a(\hat{a}_1, \cdots, \hat{a}_I)$ .

Here, each agent  $i$ 's utility is denoted as  $u_i(g_a(\hat{a}_1, \cdots, \hat{a}_I), \theta_i)$ .

**Note 4:** Generally speaking, each agent's private type is his privacy and hence has a positive value to him. Consider Step 1 in Definition 8, each agent  $i$  reports a type, either honestly or dishonestly. Note that choosing to be honest or dishonest is each agent's independent and private choice, and cannot be controlled by the designer. Hence, *from each agent's perspective, if the utility of truth-telling is not larger than but only equal to the utility of false-telling, then any reasonable agent will prefer false-telling*, because false-telling always

<sup>4</sup> For the case of action-format strategies, the designer *observes* the performance of each agent's action. As a comparison, for the case of message-format strategies, the designer *receives* each agent's message.

<sup>5</sup> If in the mechanism  $\Gamma_a$  some agent  $i$  only declares a message about how to perform an action but does not realistically perform it, then this declaration is meaningless.

hides his privacy, which will be attractive to him.  $\square$

**Note 5:** Consider Step 3 in Definition 8, after receiving the designer's suggestion, each agent performs an action either obediently or disobediently. Note that choosing to be obedient or disobedient is each agent's independent and open choice. Although each agent's action can be observed by the designer, she cannot punish any disobedient agent in Step 4 since the designer has no such power in the framework of mechanism design theory.  $\square$

### 3 Main results

Consider the multistage direct game  $\bar{G}_a$  induced by the direct mechanism  $\bar{\Gamma}_a$  described in Definition 8. According to Myerson [2], the strategy “*honest and obedient*” is the Bayesian Nash equilibrium of the game  $\bar{G}_a$ : *i.e.*, each agent  $i$  not only honestly discloses his private type in Step 1 (*i.e.*,  $\hat{\theta}_i = \theta_i$ ), but also obeys the designer's suggestion in Step 3 (*i.e.*,  $\hat{a}_i = a_i^m(\theta_i)$ ). However, in this section we will point out that Myerson's conclusion will not hold when each agent's strategy is of an action format.

**Proposition 1:** For any given social choice function  $f(\cdot) : \Theta \rightarrow X$ , suppose there is a mechanism that implements it in Bayesian Nash equilibrium, in which each agent's strategy is of an action format. Then  $f$  will not be truthfully implementable, *i.e.*, in the multistage direct game induced by the corresponding direct mechanism, “*honest and obedient*” will no longer be the equilibrium.

**Proof:** Suppose the mechanism  $\Gamma_a = (A_1, \dots, A_I, g_a(\cdot))$  implements the social choice function  $f(\cdot) : \Theta \rightarrow X$  in Bayesian Nash equilibrium, in which each agent's strategy is of an action format. By Definition 6 it induces a one-stage game  $G_a$ , the equilibrium of which is denoted as  $a^*(\cdot) = (a_1^*(\cdot), \dots, a_I^*(\cdot))$ , and  $g_a(a^*(\theta)) = f(\theta)$  for all  $\theta \in \Theta$ . Thus, there is a corresponding direct mechanism  $\bar{\Gamma}_a = (\Theta_1, \dots, \Theta_I, g_a(a^*(\cdot)))$  given by Definition 7 and a direct game  $\bar{G}_a$  given by Definition 8. In Step 1 of the direct game  $\bar{G}_a$ , there are different cases for agents  $1, \dots, I$ .

*Case 1: Each agent chooses to be honest*

Consider each agent  $i$  chooses to be “*honest*” in Step 1 of  $\bar{G}_a$  (*i.e.*,  $\hat{\theta}_i = \theta_i$ ), then in Step 2 of  $\bar{G}_a$ , the suggestion to each agent  $i$  will be  $a_i^m(\theta_i)$ . Since the equilibrium of  $G_a$  is  $a^*(\cdot) = (a_1^*(\cdot), \dots, a_I^*(\cdot))$ , then the optimal choice of each agent  $i$  in Step 3 of  $\bar{G}_a$  is to be “*obedient*” (*i.e.*, obeying the suggestion  $a_i^m(\theta_i)$  and performing the action  $a_i^*(\theta_i)$ ), given that the others also choose to be “*obedient*”. In Step 4 of  $\bar{G}_a$ , the final outcome will be  $g_a(a^*(\theta))$ , which is equal to  $f(\theta)$  for all  $\theta \in \Theta$ .

*Case 2: One agent chooses to be dishonest and the others choose to be honest*  
 Consider in Step 1 of  $\bar{G}_a$  there is one agent  $i$  that wants to protect his privacy and chooses to be “*dishonest*”, and the others still choose to be “*honest*” (*i.e.*,  $\hat{\theta}_i \neq \theta_i$ ,  $\hat{\theta}_{-i} = \theta_{-i}$ ). Then in Step 2 of  $\bar{G}_a$ , the suggestion to agent  $i$  will be  $a_i^m(\hat{\theta}_i) \neq a_i^m(\theta_i)$ , and the suggestions to the others will still be  $a_{-i}^m(\theta_{-i})$ . Since the equilibrium of  $G_a$  is  $a^*(\cdot) = (a_1^*(\cdot), \dots, a_I^*(\cdot))$ , then the optimal choice for agent  $i$  in Step 3 of  $\bar{G}_a$  will be “*disobedient*” (*i.e.*, not obeying the suggestion  $a_i^m(\hat{\theta}_i)$  but still performing  $a_i^*(\theta_i)$ ). And the optimal choices for other agents will still be “*obedient*” (*i.e.*, performing  $a_{-i}^*(\theta_{-i})$ ). By Note 5, although in Step 4 of  $\bar{G}_a$  the designer can find agent  $i$  is disobedient, she cannot punish him. In the end, the direct mechanism  $\bar{\Gamma}_a$  will yield the outcome  $g_a(a^*(\theta))$ , which is equal to  $f(\theta)$  for all  $\theta \in \Theta$ .

It can be seen from Case 1 and Case 2 that when agents  $1, \dots, i-1, i+1, \dots, I$  choose to be *honest*, then no matter whether agent  $i$  chooses “*honest*” or “*dishonest*” in Step 1 of  $\bar{G}_a$ , the optimal action for agent  $i$  in Step 3 of  $\bar{G}_a$  will always be  $a_i^*(\theta_i)$ . Obviously, from agent  $i$ ’s viewpoint, “*dishonest*” chosen in Case 2 is better than “*honest*” chosen in Case 1, because he can obtain the same outcome  $g_a(a^*(\theta))$  and at the same time hide his private type  $\theta_i$ . Furthermore, Case 2 can be generalized to everyone as follows.

*Case 3: Each agent chooses to be dishonest*

Consider each agent  $i$  chooses to be “*dishonest*” in Step 1 of  $\bar{G}_a$  (*i.e.* reporting a false type  $\hat{\theta}_i \neq \theta_i$ ), and then chooses to be “*disobedient*” in Step 3 of  $\bar{G}_a$  (*i.e.*, not obeying the designer’s suggestion  $a_i^m(\hat{\theta}_i)$  but still performing  $a_i^*(\theta_i)$ ). Then, the final outcome will still be  $g_a(a^*(\theta))$ , which is equal to  $f(\theta)$  for all  $\theta \in \Theta$ . As a result,  $f(\cdot)$  is implemented by the constructed direct mechanism  $\bar{\Gamma}_a = (\Theta_1, \dots, \Theta_I, g_a(a^*(\cdot)))$ , *not truthfully but dishonestly*. Obviously, each agent’s private type is hidden. Hence, from each agent’s perspective, “*dishonest and disobedient*” chosen in Case 3 is strictly better than “*honest and obedient*” chosen in Case 1.

To sum up, when the format of each agent’s strategy is an action, “*honest and obedient*” will no longer be the equilibrium strategy of the corresponding direct mechanism  $\bar{\Gamma}_a = (\Theta_1, \dots, \Theta_I, g_a(a^*(\cdot)))$ . Hence the revelation principle does not hold when each agent’s strategy is action-format.  $\square$

## 4 Conclusions

In this paper, we propose that in any game there are two formats of strategy (*i.e.*, an informational message or a realistic action), and the action format is common in many practical cases. In Section 2.1 we point out that the revelation principle holds when each agent’s strategy is of message for-



mat. However, when each agent's strategy is of an action format, "*honest and obedient*" will no longer be the Bayesian Nash equilibrium in the multistage direct game induced by the corresponding direct mechanism. Therefore, the revelation principle fails when each agent's strategy is of action format.

## Appendix: Notations and Definitions

Let us consider a setting with one designer and  $I$  agents indexed by  $i = 1, \dots, I$ . Each agent  $i$  privately observes his *type*  $\theta_i$  that determines his preference over elements in an outcome set  $X$ . The set of possible types for agent  $i$  is denoted as  $\Theta_i$ . The vector of agents' types  $\theta = (\theta_1, \dots, \theta_I)$  is drawn from set  $\Theta = \Theta_1 \times \dots \times \Theta_I$  according to probability density  $\phi(\cdot)$ , and each agent  $i$ 's *utility function* over the outcome  $x \in X$  given his type  $\theta_i$  is  $u_i(x, \theta_i) \in \mathbb{R}$ .

**Definition 23.B.1** A *social choice function* (SCF) is a function  $f : \Theta_1 \times \dots \times \Theta_I \rightarrow X$  that, for each possible profile of the agents' types  $\theta_1, \dots, \theta_I$ , assigns a collective choice  $f(\theta_1, \dots, \theta_I) \in X$ .

**Definition 23.B.3** A *mechanism*  $\Gamma = (S_1, \dots, S_I, g(\cdot))$  is a collection of  $I$  strategy sets  $S_1, \dots, S_I$  and an outcome function  $g : S_1 \times \dots \times S_I \rightarrow X$ .

The mechanism combined with possible types  $(\Theta_1, \dots, \Theta_I)$ , the probability density  $\phi(\cdot)$  over the possible realizations of  $\theta \in \Theta_1 \times \dots \times \Theta_I$ , and utility functions  $(u_1, \dots, u_I)$  defines a Bayesian game of incomplete information. The strategy function of each agent  $i$  in the game induced by  $\Gamma$  is a private function  $s_i(\theta_i) : \Theta_i \rightarrow S_i$ . Each strategy set  $S_i$  contains agent  $i$ 's possible strategies. The outcome function  $g(\cdot)$  describes the rule for how agents' strategies get turned into a social choice.

**Definition 23.B.5** A *direct mechanism* is a mechanism  $\bar{\Gamma} = (\bar{S}_1, \dots, \bar{S}_I, \bar{g}(\cdot))$  in which  $\bar{S}_i = \Theta_i$  for all  $i$  and  $\bar{g}(\theta) = f(\theta)$  for all  $\theta \in \Theta$ .<sup>6</sup>

**Definition 23.D.1** A strategy profile  $s^*(\cdot) = (s_1^*(\cdot), \dots, s_I^*(\cdot))$  is a *Bayesian Nash equilibrium* of mechanism  $\Gamma = (S_1, \dots, S_I, g(\cdot))$  if, for all  $i$  and all  $\theta_i \in \Theta_i$ ,  $\hat{s}_i \in S_i$ , there exists

$$E_{\theta_{-i}}[u_i(g(s_i^*(\theta_i), s_{-i}^*(\theta_{-i})), \theta_i) | \theta_i] \geq E_{\theta_{-i}}[u_i(g(\hat{s}_i, s_{-i}^*(\theta_{-i})), \theta_i) | \theta_i].$$

**Definition 23.D.2** The mechanism  $\Gamma = (S_1, \dots, S_I, g(\cdot))$  *implements the social choice function*  $f(\cdot)$  *in Bayesian Nash equilibrium* if there is a Bayesian

<sup>6</sup> The bar symbol is used to distinguish the direct mechanism from the indirect mechanism.

Nash equilibrium of  $\Gamma$ ,  $s^*(\cdot) = (s_1^*(\cdot), \dots, s_I^*(\cdot))$ , such that  $g(s^*(\theta)) = f(\theta)$  for all  $\theta \in \Theta$ .

**Definition 23.D.3** The social choice function  $f(\cdot)$  is *truthfully implementable in Bayesian Nash equilibrium* (or *Bayesian incentive compatible*) if  $\bar{s}_i^*(\theta_i) = \theta_i$  (for all  $\theta_i \in \Theta_i$  and  $i = 1, \dots, I$ ) is a Bayesian Nash equilibrium of the direct mechanism  $\bar{\Gamma} = (\Theta_1, \dots, \Theta_I, f(\cdot))$ . That is, if for all  $i = 1, \dots, I$  and all  $\theta_i \in \Theta_i$ ,  $\hat{\theta}_i \in \Theta_i$ ,

$$E_{\theta_{-i}}[u_i(f(\theta_i, \theta_{-i}), \theta_i) | \theta_i] \geq E_{\theta_{-i}}[u_i(f(\hat{\theta}_i, \theta_{-i}), \theta_i) | \theta_i].$$

**Proposition 23.D.1** [1]: (*The Revelation Principle for Bayesian Nash Equilibrium*) Suppose that there exists a mechanism  $\Gamma = (S_1, \dots, S_I, g(\cdot))$  that implements the social choice function  $f(\cdot)$  in Bayesian Nash equilibrium. Then  $f(\cdot)$  is truthfully implementable in Bayesian Nash equilibrium.

## Acknowledgment

This paper has been conceived for more than ten years. In order to memorize this invaluable period, here I would like to cite a famous verse written by Ting-Jian Huang in China Song Dynasty, “*Tao Li Chun Feng Yi Bei Jiu, Jiang Hu Ye Yu Shi Nian Deng*”. Thanks a lot to my love: Fang, Yue, Xing, Chen.

## References

- [1] A. Mas-Colell, M.D. Whinston and J.R. Green, Microeconomic Theory. Oxford University Press, 1995.
- [2] R. Myerson, Optimal coordination mechanisms in generalized principal-agent problems. Journal of Mathematical Economics, vol.10, 67-81, 1982.
- [3] Y. Narahari et al, Game Theoretic Problems in Network Economics and Mechanism Design Solutions. Springer, 2009.
- [4] H. Bester and R. Strausz, Contracting with imperfect commitment and the revelation principle: The single agent case. Econometrica, Vol.69, No.4, 1077-1098, 2001.
- [5] Epstein and Peters, A revelation principle for competing mechanisms. Journal of Economic Theory, 88, 119-160 (1999).
- [6] A. Kephart and V. Conitzer, The revelation principle for mechanism design with reporting cost. In Proceedings of the ACM Conference on Electronic Commerce (EC), Maastricht, The Netherlands, 2016.